



King's Research Portal

DOI:

[10.1016/j.jhep.2017.08.007](https://doi.org/10.1016/j.jhep.2017.08.007)

Document Version

Peer reviewed version

[Link to publication record in King's Research Portal](#)

Citation for published version (APA):

Thompson, R. J. (2017). Sequencing of transporter genes in cholestasis: we are still learning. *Journal of Hepatology*. <https://doi.org/10.1016/j.jhep.2017.08.007>

Citing this paper

Please note that where the full-text provided on King's Research Portal is the Author Accepted Manuscript or Post-Print version this may differ from the final Published version. If citing, it is advised that you check and use the publisher's definitive version for pagination, volume/issue, and date of publication details. And where the final published version is provided on the Research Portal, if citing you are again advised to check the publisher's website for any subsequent corrections.

General rights

Copyright and moral rights for the publications made accessible in the Research Portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognize and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the Research Portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the Research Portal

Take down policy

If you believe that this document breaches copyright please contact librarypure@kcl.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.

Accepted Manuscript

Sequencing of transporter genes in cholestasis: we are still learning

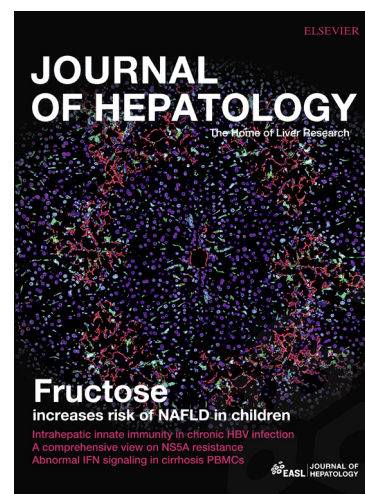
Richard J. Thompson

PII: S0168-8278(17)32213-4
DOI: <http://dx.doi.org/10.1016/j.jhep.2017.08.007>
Reference: JHEPAT 6633

To appear in: *Journal of Hepatology*

Received Date: 10 August 2017
Accepted Date: 15 August 2017

Please cite this article as: Thompson, R.J., Sequencing of transporter genes in cholestasis: we are still learning, *Journal of Hepatology* (2017), doi: <http://dx.doi.org/10.1016/j.jhep.2017.08.007>



This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Sequencing of transporter genes in cholestasis: we are still learning

Richard J Thompson

Institute of Liver Studies
King's College London
London
UK

Correspondence:

Postal Institute of Liver Studies
 King's College Hospital
 Denmark Hill
 London
 SE5 9RS
 UK
Telephone +442032994296
Email richard.j.thompson@kcl.ac.uk

Grant support

RJT received support from NIH R01 DK094828

Conflicts of interest

RJT received support from Shire, Albireo, Alexion, Arcturus, GSK and Retrophin. None are however relevant to the content of this manuscript.

Sequencing of DNA has never been easier. It is now perfectly possible to sequence an individual's entire genome from a few micrograms of DNA, obtained from peripheral blood or elsewhere. The cost of sequencing has also fallen dramatically. The biggest hurdle is now in making sense of the data. Variation in DNA sequence contributes a large part of the phenotypic differences between individuals. However most of the variation occurs in non-coding DNA; the overwhelming majority of which we cannot yet interpret. Variation within genes is much better understood; but even here not everything is as simple as it might seem.

"Simple" genetics might consider highly penetrant alleles, causing Mendelian phenotypes, be they dominant or recessive. That is to say, that those with a given genotype have highly predictable manifestations. At the other end of the spectrum are relatively common variants that have been associated with particular diseases. In the latter case many unaffected individuals will have the risk allele; although it is more common in those with the disease phenotype. In between these extremes are variants which are identified as making intermediate contributions to the phenotype in question. This latter group might be considered "contributory", "predisposing" or even "modifiers" of a condition. In truth however, many variants contribute to every disease, it is just the extent of their contribution that varies. There are probably no fully penetrant alleles, or indeed any discrete categories of contribution. Worse still, because of the effect of other variants, and non-genetic factors, one variant might contribute more in one individual than in another.

All of this makes the interpretation of genetic variants surprisingly complex. There has been a rapid growth in websites and software designed to help us interpret genetic variation. There have been guidelines, most notably from the American College of Medical Genetics and Genomics (ACMG) (1), to help us. The software tools are only as good as the information available to them; this is mostly from the public domain, though some have their own databases. The ACMG guidelines are now widely employed; though they probably increase consistency, more than improving accuracy. To date all these tools have still been directed at highly penetrant or Mendelian models of disease.

We understand variation in some genes better than others. The liver disease gene with the best understood variation is probably HFE, underlying genetic haemochromatosis. The contributions to disease of the three major variants have been very well characterised(2). The risks associated with all combinations of these alleles has been defined. However this doesn't predict those who will get haemochromatosis; only those with increased risk. In more recent times a common variant in PNPLA3 has been associated with both NAFLD and alcohol-associated liver disease. However in this case the penetrance and the contribution to disease are so low that it is of little use in making a diagnosis, though it may be of use in predicting response to treatment (3).

A growing list of genes have been associated with cholestatic liver disease. It is now 19 years since the first three were described (4-6). This means that considerable numbers of patients have been sequenced for these genes, and some prediction of the contribution of genetic variants can be made. Merely collecting genetic data is of little value however, it is only when this is combined with high quality phenotypic data can we begin to learn. In this issue Keitel and colleagues (7) describe 154 variations in *ATP8B1*, *ABCB11* and *ABCB4* in 427 individuals, all suspected of having genetic cholestasis. In this cohort of patients several previously described disease-causing variants were found. Several novel putatively disease-causing variants were identified. These are important additions to the ever-growing catalogue of variants for which there is good evidence of major pathological effect. Protein truncating mutations form the minority, though their pathogenicity is generally clear. Consensus splice site mutations, affecting the first or last 2 bases of an intron are also generally easy to classify as detrimental. The current difficulty really lies with missense changes. Such variants are frequently thought of as changes in the amino acid sequences. In truth they are alterations in genomic DNA. If transcription occurs normally they become changes in pre-mRNA. As such they may affect splicing and processing into mRNA. Only if these steps are overcome can such changes become variants in the polypeptide (8).

The molecular consequence of variants in the amino acid sequence of proteins can be analysed in several ways. The most powerful evidence in favour of pathogenicity is the observation of such a change, in trans, with another known pathogenic change. All other interpretations of function are open to criticism. Commonly applied observations include the degree of evolutionary conservation of the same protein, and between similar proteins, at both the nucleotide and amino acid level. The functional change predicted by in silico models can be helpful, but the tools available still lack enough data. In vitro models can be used to look at all the stages of mRNA and protein production, and function. However no in vitro experiment is perfect, yet! Keitel and colleagues combine the widely available

predictive tools, along with less used computer modelling. Future patients will tell us which have worked best, and the determined pathogenicity of known variants will increase or decrease accordingly.

Previous studies have shown that common variants in these genes are over represented in patients with “acquired” cholestasis. Of course such patients probably harbour variants which impair the quantity or function of the gene/protein to a lesser degree than the better described “disease-causing” mutations. The idea that “disease-predisposing” or “modifier” variants are in fact similar is highlighted by data in the current paper. Some common variants appear to be more common than is seen in the general population; in patients with no other mutations, where they would be predisposing or in patients with pathogenic mutations, where they would be modifiers. In all experiments examining allele frequencies the biggest problem is control data. The bigger, and better matched, the control set is the more credible are any differences in frequency between them and patients. The data in Keitel’s paper compare the observed frequencies with those seen in the Gnomad database (9). This is currently the largest publically-available set of allele frequencies. It is in fact an aggregation of data from other projects, and includes information from 138,000 individuals. The presence of a variant in such a large dataset does not exclude it from being highly pathogenic, as many are very rare. On the other hand many variants have very different frequencies in different populations. The data presented in Gnomad are broken down into 7 large groups and “others”. Brief inspection will show that some populations have been much more extensively investigated than others; e.g. Ashkenazi Jews and Finns. Europeans make up almost half the chromosomes sequenced, though we do now have much better data on individuals from other regions of the world than we had even one year ago. Keitel compares the overserved frequencies with those seen in the entire Gnomad dataset, and with those seen in Europeans, as that was thought to be the origin of most of their patients. As the quantity and quality of control data grows this and previous analyses, based on comparison with controls, will need re-examination.

DNA sequencing is now increasingly cheap and easy. Data analysis however remains complex. Variants in genes should be thought of as contributing to the phenotype; the most difficult decision in to what extent. The sequencing being carried out today, be it in patients and controls, will help build the datasets that will inform future research and, more importantly, diagnosis.

References

1. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, Grody WW, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med* 2015;17:405-424.
2. Powell LW, Seckington RC, Deugnier Y. Haemochromatosis. *Lancet* 2016;388:706-716.
3. Anstee QM, Seth D, Day CP. Genetic Factors That Affect Risk of Alcoholic and Nonalcoholic Fatty Liver Disease. *Gastroenterology* 2016;150:1728-1744 e1727.

4. Bull LN, van Eijk MJ, Pawlikowska L, DeYoung JA, Juijn JA, Liao M, Klomp LW, et al. A gene encoding a P-type ATPase mutated in two forms of hereditary cholestasis. *Nat Genet* 1998;18:219-224.
5. de Vree JM, Jacquemin E, Sturm E, Cresteil D, Bosma PJ, Aten J, Deleuze JF, et al. Mutations in the MDR3 gene cause progressive familial intrahepatic cholestasis. *Proc Natl Acad Sci U S A* 1998;95:282-287.
6. Strautnieks SS, Bull LN, Knisely AS, Kocoshis SA, Dahl N, Arnell H, Sokal E, et al. A gene encoding a liver-specific ABC transporter is mutated in progressive familial intrahepatic cholestasis. *Nat Genet* 1998;20:233-238.
7. Droge C, Bonus M, Baumann U, Klindt C, Lainka E, Kathemann S, Brinkert F, et al. Sequencing of FIC1, BSEP and MDR3 in a large cohort of patients with cholestasis revealed a high number of different genetic variants. *J Hepatol* 2017.
8. Byrne JA, Strautnieks SS, Ihrke G, Pagani F, Knisely AS, Linton KJ, Mieli-Vergani G, et al. Missense mutations and single nucleotide polymorphisms in ABCB11 impair bile salt export pump processing and function or disrupt pre-messenger RNA splicing. *Hepatology* 2009;49:553-567.
9. Gnomad Browser <http://gnomad.broadinstitute.org>